

A First Step toward Recommendations Based on the Memory of Users

Florian Marchal
LORIA - Université de Lorraine
Campus Scientifique, B.P. 239
54506 Vandœuvre - France
florian.marchal@loria.fr

Sylvain Castagnos
LORIA - Université de Lorraine
Campus Scientifique, B.P. 239
54506 Vandœuvre - France
sylvain.castagnos@loria.fr

Anne Boyer
LORIA - Université de Lorraine
Campus Scientifique, B.P. 239
54506 Vandœuvre - France
anne.boyer@loria.fr

Abstract—Most of recommender systems build their predictions by analysing the preferences of users. However, there are many situations, such as in intelligent tutoring systems, where recommendations of pedagogical resources should rather be based on their memory. So as to infer in real time and with low involvement what has been memorized by users, we highlight in the paper the link between gaze features and visual memory. We designed a user experiment where different subjects had to remember a large set of images. In the meantime, we collected about 19,000 fixation points. Among other metrics, our results show a strong correlation between the relative path angles and the memorized items. It is thus possible to predict the users' memory status by analyzing their gaze data while interacting with the system, so as to provide recommendations that fits their learning curve.

Index Terms—Learner modeling; Gaze data; Eye-Tracking; Visual memory; Recall

I. INTRODUCTION

Memory and Learning are closely related, since it is often necessary to memorize notions or methods in order to learn concepts or acquire competences. Most of the time, evaluating the students' level of attainment in schools or within intelligent tutoring systems relies on exercises, tests or projects that only allow a partial validation of declarative knowledge. Indeed, the proposed tests, whether anticipated or terminal, only cover a small subset of the points addressed by the educational program. It is therefore difficult to have a comprehensive view of the students' level. Furthermore, these tests do not completely determine the optimal character of learning concepts, mainly because they allow to evaluate the answers of the students, but not necessarily the process that led to these results.

The development of new technology now makes possible the recording of gaze data (fixation points, durations, saccades, areas of interest, scan path) when interacting with a computer system, or even when reading or writing on a piece of paper. This technology, is currently very expensive, but the gradual integration of eye-trackers in the hardware devices by manufacturers (computers, tablets, smartphones) will make it more accessible within the next five years.

In this paper, we aim at predicting the memory status of students, throughout the process of acquiring knowledge. This work has two long-term goals. First, we could provide a precise feedback to the teachers about the progress made by

the students (what has been learned, understood or memorized by each of them). Second, we can consider providing personalized recommendations of educational resources to students, based on their progress. A recent study on Massively Open Online Courses (MOOCs)¹ shows that only 6.5% of users really complete them. The length of the course appears to be the critical factor, since it is the same for every user. In this context, recommending resources that fit the active user's learning curve, in addition to their interests, may increase their satisfaction and the completion rate of the courses. Three steps are required in recommender systems: (1) modeling users (in our case, we aim at modeling users' memory), (2) computing adapted recommendations based on users' model, and (3) providing these recommendations in the good manner at the right time for an efficient metacognitive regulation (planning, monitoring, evaluating). We focus here on the first step.

Modeling the memory of users in a traditional way can be very time-consuming and requires a high user involvement. Thus, we investigate the possibility to infer the memorized items through gaze data. Within the frame of recommenders, eye-tracking systems are mainly used to infer user preferences [1], to evaluate the accuracy of a recommender system [2], or to evaluate users' learning strategy [3]. Some works investigated the link between gaze data and memory, but most of them focus on the recognition process [4]. We propose here to use gaze features to model what has been recalled by users.

Section II offers an overview of the literature as regards memory and eye-tracking usages to model cognitive processes. Section III is dedicated to the presentation of our experiment. Section IV presents the analysis of the data and the results.

II. RELATED WORK

The related work presented in this article is related to the modeling of phenomena connected with human memory. We are particularly interested in eye behavior to study the memorization process. Section II-A introduced the literature about the functioning of memory and memory processes. Section II-B focuses on the analysis of the gaze behavior.

¹Katy Jordan, University of Texas, 2013 may
<http://techcrunch.com/2014/03/03/study-massive-online-courses-enroll-an-average-of-43000-students-10-completion/>

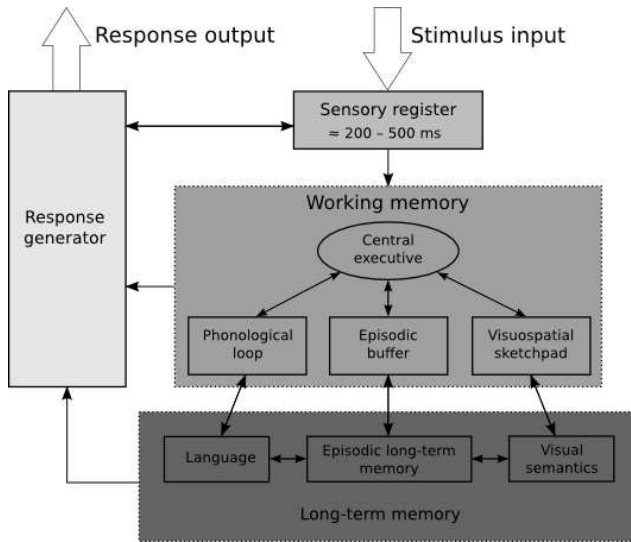


Fig. 1. Model of short term and long term memories.

Finally, we explain in the section II-C how we will analyze the gaze behavior to model the memorization process.

A. Memory

Memory and memorization processes have been, and still are, a source of debate and research in cognitive psychology.

The Atkinson-Shiffrin model, defined by [5], is the most widely accepted model to represent the memory. As shown in Figure 1, it consists of three distinct memory modules that interact: the Sensory Register (SR), which retains the sensory impressions when the stimulus stops; the Short-Term Store (STS), storage space which interacts with the Long-Term Store (LTS).

However, this model has been re-designed by [6]. The STS which was described as a unique bridge between the SR and the LTS has been replaced by the working memory. The latter is composed of three components (the phonological loop, the visuospatial sketchpad, and the episodic buffer) subordinated to a fourth component (the central administrator). These three components have their role and their own way of encoding information.

To investigate this working memory, one must know its limits. The memory span is the maximum number of simultaneous storable items in this memory. Miller [7] already explained in 1956 that this span is 7 ± 2 items. Despite this limitation, the working memory is able to group common elements into a complex item in order to preserve it. This operation, called chunking, has been studied by Gobet and Simon [8] within the frame of connection failures. They found that a good usage of memory consisted in making a large collection of assemblies of pieces, which resulted from the chunking, move from the working memory to the long-term memory, and vice versa.

We will not consider directly the working memory but the memory processes that pass through it. Jonassen and Tesser [9] clearly marked the distinction between two different processes of memory access:

- the **recognition** is the mental process which consists in comparing and associating a current and past stimuli. In other words, this is the ability to recognize previously encountered items (objects, people, and so on), when they are represented;
- the **recall** involves to remember a stimulus which is not physically present.

Brain imaging studies [10], [11] show that these two processes have separated and distinct neuronal connections. Clariana and Lee [12] also show in their study that the combined use of recognition and recall tasks allow a better learning gain, rather than using them separately. This would tend to show that these two processes complete each other, more than they interfere.

B. Gaze Analysis

The current use of eye-tracking systems aim at understanding the user's behavior. As an example, Chen and Pu [2] use it successfully to assess the effectiveness of their recommender system. In the context of web platform, Sari *et al.* [13] address measures to be applied when one wishes to analyze a user experience through eye-tracking techniques. But other works go further and show that we can distinguish the indices of cognitive mechanisms in the gaze behavior. Lu and Jia [1] infer users' preferences from gaze data with a good efficiency for the use of a web platform. Kanan *et al.* [14] even implement what had been hypothesized by Yarbus [15]: the task performed by a user could be predicted from the analysis of the way his eyes behave while doing this task.

In the field of cognitive neuroscience, several studies address the recognition process by studying it through eye-tracking techniques. Hannula [16] proposes a quite complete overview of recognition, by highlighting the interest of this method with particular populations (young children, elderly people, ...) and the ability to compare humans and animals through this. She also shows that the combination of digital imaging and eye-tracking techniques allows a better understanding of the link between the brain and users' behavior. More recently, Hannula *et al.* [4] conducted a study on the recognition process, where the stimuli proposed to users are faces of people. The names of these people are previously presented to participants, jointly with a familiarity questionnaire, so as to distinguish those who are considered as "famous" from those who are "unknown" for the user.

Even more recently, in the field of image study, some authors became interested in the concept of memorability, through the use of eye-tracking systems. Bylinskii *et al.* [17] use memorability and gaze behavior of users to show that one can reasonably predict whether an image will be subsequently recognized or not. Borkin *et al.* [18] also studied the memorability of charts, using the eye-tracking to monitor how users recognize them. However, it is important to understand that an item that is easily memorable and recognizable will not necessarily be recalled by the active user during a memory test. The memorability only represents the chances that a large population has to remember an item. It consequently requires

a large number of users to be computed for each item, and this is not what we want to achieve. On the long-term, we aim at predicting recalled items, whatever these items are (i.e. without pre-computing a memorability score), by only relying on the active user's gaze behavior.

Closer to what we want to accomplish, Bondareva *et al.* [3] use the gaze data of users while they interact with an intelligent tutoring system. They try to predict in real time the efficiency of the users' learning process, thus modeling the users' effectiveness during their interactions. Steichen *et al.* [19] push the analysis further by studying, within the frame of a work on graphical displays, the gaze data but also the transition sequences between the areas on which the eyes of users arose. This is still an on-going research, but they hypothesize that some sequences may reflect certain characteristics of the users as auditory working memory, visual working memory or perceptual speed. Other sequences reveal the difficulty of the task.

C. Discussion

As stated before, our goal consists in modeling the memorization process of users so as to better assist them in e-learning environments. More precisely, we carry out an on-going research on the way to predict what is recalled by users from gaze features. We are thus furthering the works of Bondareva, Steichen *et al.* [3], [19]. Indeed, while most of works aim at understanding users' behavior from gaze data [13], Steichen *et al.* [19] followed eye movements to infer users' characteristics, such as perceptual speed and working memory. Nevertheless, we wish to establish a broader context than graphical displays or tutorial systems. The works of Bondareva *et al.* [3] and Steichen *et al.* [19] involve the creation of predefined Areas Of Interest (AOIs), which can hardly be applied generally. Bylinskii *et al.* [17] and Borkin *et al.* [18] are very close to what we want to achieve in the medium term, but they focus their approach on the nature of the images and displays, while we would like to model the behavior and capabilities of users on any tasks requiring an effort of memory. We expect to extend previous works by finding correlations between gaze features and memorized items, without the use of predefined AOIs. Let also notice that the work of Hannula *et al.* [16], [4] is mainly oriented on the recognition, and does not aspire to memory prediction. However, they propose interesting ways to extract the gaze features that could be used to model the memory abilities of users.

III. EXPERIMENT

So as to address our objective which consists in finding a way to model user memory from gaze data, i.e. identifying gaze features that are the best candidates to infer memory, we conducted a user study where participants were invited to complete a memory test with images. During this test, we collected gaze data so as to compare these pieces of data with users' answers and emphasize statistical links between them.

A. Hypotheses

Our main hypothesis relies on the fact that the gaze behavior of a participant can be decomposed in several gaze features, assuming that some of these features could be useful to predict the current memory state of this user. This leads us to our hypothesis **H0**.

H0. The analysis of gaze features can reveal which items are recalled, and which are not, during human-computer interactions.

If the statistical link between these gaze features and the results of the memory test is confirmed, this work will be a first step toward identifying the features that are the best predictors of the memory state. Thanks to this analysis of features, we could provide recommendations of items based on and adapted to user memory. This hypothesis seemed reasonable as regards the results obtained for the recognition of faces [4], and work on the prediction of the quality of the learning process from gaze data [3].

We also have two hypotheses in relation with our experiment:

H1. The different categories of images provided to users during the memory test will not influence the link between the gaze behavior and the capability to recall these images.

H2. The age of participants does not affect the link between the gaze behavior and the capability to recall the images.

B. Material

A Tobii X1 Light Eye Tracker was used to monitor the subjects' eye movements and the software Tobii Studio was used to record the gaze data and transform them into fixations. Besides, so as to define a task that focuses on the recall process, we defined a set of 1,395 images depicting real-world objects. These images are provided by Brady *et al.* [20]. We chose 93 categories among the 200 available categories. Each category was made of 15 different images.

Konkle *et al.* [21] used the same corpus of images, and asked participants to rate every category according to their colors, their form, their genre and the overall perceptual of each of its elements. We re-use these pieces of information coming from real users to select the categories which offer a strong salience (i.e. a high overall perceptual level). In this way, even if several images belong to the same category, they are easily identifiable.

The study of Konkle *et al.* [21] shows that the overall perceptual level within a category has no influence on the memorization of items that compose this category. However, this criteria was useful within the context of our experiment. As our participants had to describe the images they saw, we wanted to be sure they could distinguish them.



Fig. 2. Examples of category used.

C. Participants

We recruited 24 french users for the need of our experiment. They were contacted by email. We conducted our experiment over a period of 2 weeks. Our population were composed of 14 men and 10 women. They were between 23 and 62 years old (mean = 32 years old). All participants had normal or corrected-to-normal visual acuity. This study being exploratory, and the first step consisting in collecting results from a control population (before focusing on students in an intelligent tutoring system), we chose an age spectrum as large as possible from an academic population, thus accustomed to using their memory.

The global duration of the experiment was quite variable from a user to another. Indeed, we wanted each participant to take the time to recall as many items as they could with no limit of time. Thus, the time needed to complete the test phase is comprise between 10 and 20 minutes according to the participants. The duration of the learning phase (the phase during which the gaze behavior was recorded) remains the same for all participants.

D. Procedure

In this section, we present the experiment we developed to validate our assumptions. We took inspiration from Maxcey and Woodman [22] to define our experiment protocol. Our experiment is made of two phases, with a short break between them. For each participant, we randomly chose 12 categories, and 6 images per category. Thus, each user was asked to recall 72 images.

Learning phase: During this first step, the randomly chosen images were presented one after another in a random order (cf. Figure 3). Participants were instructed to study details on each picture for a later memory test. They were told that the test would require very detailed information, and that remembering only the category would not help. Each image was displayed during 5 seconds, interleaved by a 500ms center fixation cross. Each image has been configured to be displayed in a square of 500x500 pixels, in the middle of the screen, the rest of the screen being grey and neutral.

Test phase: After a short break, participants were instructed to describe on a sheet all the images they can remember from the first step, with no limit of time (cf. Figure 4). Finally, they were asked to match their answers with memorized items by browsing the 72 images on a single page. The correct images have been categorized as “remembered”, and all the others as “not remembered”. The instructions in Figures 3 and 4 were provided in French to french users, but have been translated for the need of this paper.

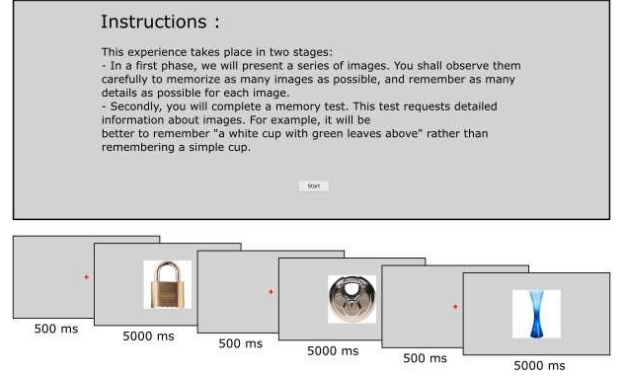


Fig. 3. Learning phase of our experiment protocol.

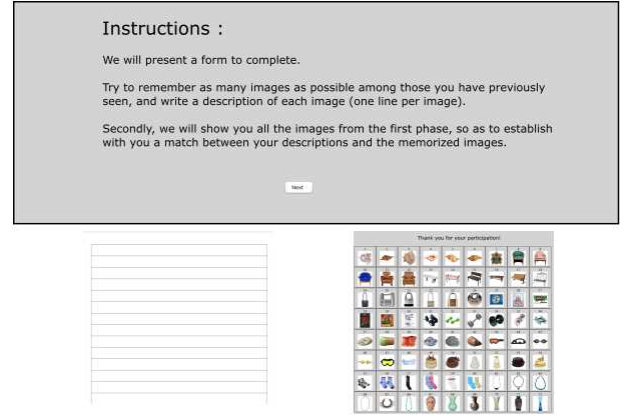


Fig. 4. Test phase of our experiment protocol.

IV. RESULTS

In order to analyze our data, we considered that each tuple $\langle participant, image \rangle$ were unique. We do not rule out the possibility that there may be classes of remarkable gaze behaviors that could fit to several participants. But initially we rely on the conclusions of Greene *et al.* [23] according to which it is possible to find the participant’s identity from his/her gaze traces. This let us foresee that each participant has a unique way of observing an image. Moreover, Kanan *et al.* [14] also explain that the analyses of Greene *et al.* [23] have failed, because they were based on all the participants. Our approach will consequently consist in analyzing the gaze behavior, participant per participant.

A. Data Processing

The processing of gaze data collected during the learning phase relies on four steps:

- 1) **Retrieving raw data:** we used the software Tobii Studio to analyze raw data and to transform them in fixation points. A data extraction program has been implemented to parse data in the files created by Tobii Studio during the experiment;

- 2) **Reconstruction of saccades:** From the data that we got from our data extraction program, another program has rebuilt saccades, thanks to geometric computations. A saccade is defined as the path crossed between two consecutive fixation points;
- 3) **Dynamic building of areas of interest:** the classical approach when one uses areas of interest consists in defining the areas *a priori*, according to the meaning we can give to them. For example, Steichen *et al.* [19] split up their charts in order to have areas of interest with the same meaning and to be able to compare them (e.g. the text, the legend, the top of the bars on an histogram, ...). Given the fact that our images are very different from each other (in particular a great variability between categories), it was not possible to manually define *a priori* AOIs, and we opted for a different approach. Althoff *et al.* [24] both use the classical approach and an approach where the AOIs are directly built from the fixation raw data of participants. They group these fixation raw data according to the distance between them. We took inspiration from this second approach and used a clustering algorithm, called DBSCAN (Density-Based Spatial Clustering of Applications with Noise) to group the fixation raw data into areas of interest. This clustering algorithm has been chosen for its ease of use, and because it allows to group fixation raw data without knowing in advance the expected number of groups. Some fixation data cannot be associated to any group, because of the low density around them. In our case, we considered that these data had no interest. Our assumption started from the statement that, if there were not enough significant gaze activity, Tobii Studio would not have created fixation points to these places. Finally, it is important to understand that each area of interest was associated to a tuple $\langle \text{participant}, \text{image} \rangle$, starting from the principle that each participant has his own way to observe an image;
- 4) **Computation of gaze features:** from all these pieces of data created in steps 1 to 3, our second program computed the gaze features and stored them in files that we can use for the statistical analysis. We chose a set of 18 gaze features (see Table II), according to the features that were defined in two user studies [16], [3].

From the data collected in steps 1 to 3, we especially measured the following gaze features (among others):

- **Number of fixations**, the number of fixations made while the participant was looking at the image. It gives an indication of the intensity with which the participant observed the image;
- **Duration of fixation** (Sum, Mean, and Standard Deviation), the amount of time the fixation lasts when the participant was looking at the image. It gives an indication of the relative importance of this fixation;
- **Distance from a fixation**, the distance between the position of the fixation and the border of the screen. It gives an indication of the distribution of fixations on the image;
- **Length of a saccade** (Sum, Mean, and Standard Deviation), the distance that we computed between two successive fixation points corresponding to the saccade (d on Figure 5). It gives an indication of the participant's gaze speed while his/her eyes scan the image ;
- **Relative angle between two saccades** (Sum, Mean, and Standard Deviation), the angle formed between two consecutive saccades (β on Figure 5). It gives an indication about the way the participant's eyes scanned the image;
- **Absolute angle of a saccade** (Sum, Mean, and Standard Deviation), the angle formed between the horizontal axis and the saccade (α on Figure 5). It gives a complementary indication about the way the participant's eyes scanned the image;
- **Number of areas of interest**, the number of areas of interest built thanks to the clustering algorithm DBSCAN. It gives an indication of the distribution of the elements that catch the gaze on the image;
- **Number of fixations in an AOI** (Mean), the number of fixations which are inside a given area of interest. It gives an indication of the size of each AOI, and the relative importance of this AOI;
- **First degree entropy**, the predictability of transitions between the location of a fixation point and the location of the previous one. It gives an indication of proportion of randomness among the transitions between the areas of interest of the participant. This entropy was calculated according to the same principle used in the analysis of Althoff *et al.* [24].

In order to compute the first degree entropy, we build a $N \times N$ transition matrix (N being the number of areas of interest). As an example, this matrix is then filled so that the transition $F_i\{ZI_2\} \rightarrow F_{i+1}\{ZI_3\}$ increases the value of the transition matrix in column 3 and row 2. Finally, we apply the following formula to determine the first degree entropy (S_1):

$$S_1 = \frac{I_{Ctot} + I_{Ltot} - I_{Cell}}{I_{Ctot}} \quad (1)$$

With the following notations:

- $I_{Cell} = \sum_{i=0}^n \sum_{j=0}^n (P_{i,j} * \log_2(\frac{1}{P_{i,j}}))$;
- $I_{Ltot} = \sum_{i=0}^n [(\sum_{j=0}^n (P_{i,j})) * \log_2(\frac{1}{\sum_{j=0}^n (P_{i,j})})]$;
- $I_{Ctot} = \sum_{j=0}^n [(\sum_{i=0}^n (P_{i,j})) * \log_2(\frac{1}{\sum_{i=0}^n (P_{i,j})})]$;
- $P_{i,j}$: the relative probability of a cell within the transition matrix.

The goal of this data processing is to identify the possible correlations between some gaze features and the fact that participants have memorized images or not. We will present the results of this statistical analysis in the next subsection.

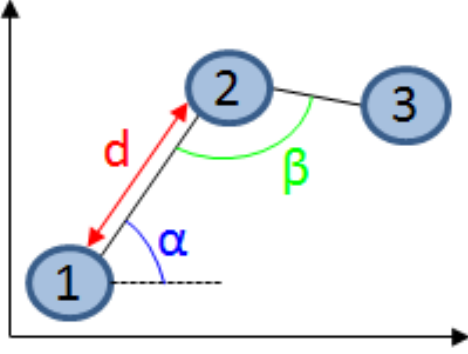


Fig. 5. Fixations, saccades, and measures.

B. Statistical Analysis

We first computed the general statistics of usages of our participants during the experiment. These statistics are shown in Table I. As expected, there were more forgotten images than remembered ones, due to a large set of images in the learning phase. We got an average of 10 fixation points per image and per user.

In a second time, we computed the 18 gaze features for each image and each participant, and checked if there is a correlation between these features and the fact that the images have been recalled by the participant or not. We first used an analysis of covariance (ANCOVA) to test the gaze data, all the images and all the participants brought together. The results were marginally significant. We thus studied the distribution of data with the Shapiro-Wilk test, and found that it does not follow a normal form, which reduces the statistical power of our test. By consequent, we chose to apply a permutation test. The latter is a non-parametric test which better reflects the distribution of our data. In Table II, we present the results of this test. It appears that the sum of the relative angles and the number of fixations are good indicators of the recalled items, with respective residual sum of squares to 0.95 and 0.7. These results are statistically significant at a 0.99 level. The third column of the Table II shows the probability that these results are due to chance.

These results prove that there is a direct link between some gaze features and the memorization of images. This validates our hypothesis H0. It also reinforces our conviction that it will be possible to predict what the users remember, by using these gaze features.

In order to study the impact of the categories of images on the memorization process, we counted the number of remembered items and forgotten items for each category, and each user. We normalized the data when summing the score for all users, since categories were randomly chosen for each user, which means that users did not necessarily look at the same categories. This score for each category can be seen as the global ease of memorization of this category for all users. Then we computed a score of ease of memorization of the proposed memory test for each user, as the sum of eases of memorization of the different categories they have been

confronted with (all the users have seen the same number of categories). These scores are shown in the second column of the Table III. The higher this score is, the easier the test was for a given user. In the third column, we also include the memorization ratio of each user, which corresponds to the number of remembered items of a given user divided by the number of displayed images during the learning phase (i.e. 72). Finally, we counted for each user the number of gaze features that are significantly correlated to the memorized items.

The distribution of data in Table III follows a normal form, and the correlation between the memorization ratio of each user and the ease of memorization of each memory test associated to a given user has been highlighted with an ANOVA ($p = 2.5e - 05$). This confirms the relevancy of our measure of ease of memorization. However, there were no correlation between the ease of memorization and the number of gaze features that can predict the memory state. This validates our hypothesis H1.

Finally, we analyzed our data by grouping users in age brackets. We did not found any evidence that the age of participants might influence the gaze behavior and our ability to predict the memory state. However, we need additional users and an equal proportion of users per age group to fully validate our hypothesis H2.

V. CONCLUSION AND FUTURE WORKS

The democratization of Information and Communication Technologies for Education (ICTE) has greatly upset the usages and modes of learning during the last decade [25]. For example, the widespread use of platforms such as Moodle in universities allowed to organize and structure the access to educational resources, and facilitated the sharing of resources between teachers [26]. However, it also highlighted the extreme complexity inherent in IT environments for human learning [27]. The first difficulty lies in the number and variety of the proposed educational resources, making access to these resources more opaque [28]. The huge size of the catalog of proposed items can be explained by the wide variety of media (computer-assisted presentations, books, videos, podcasts, websites, exercises, ...), the different public (especially their level of knowledge/skills already acquired) and the pedagogical approaches (participatory, by project, by goals, by problems, by competences, ...). This complexity leads teachers to build a multitude of *ad hoc* learning scenarios. Learners, when confronted with such systems, are often buried under the mass of information and struggle to find the right resources adapted to their skills and knowledge. This mass data problem is exacerbated with the advent in 2012 of massive open online courses (MOOCs). Each course is followed by an average of 43,000 students and there are already several thousand MOOCs, this number growing by 200% each year.

Within this context, the use of recommender systems and personalization techniques is more and more required [29]. Nevertheless the current recommender systems rely on the modeling of users' preferences, so as to suggest lessons related to their interests. As previously stated, only 6.5% of users who

TABLE I
STATISTICS OF POPULATION.

	Max	Min	Sum	Mean	Std
Remembered images	43	10	610	25.42	9.264
Forgotten images	62	29	1,118	46.58	9.264
Number of fixations per image	19	0	19,230	10.788	3.681
Fixation durations (ms)	9,476	0	8,229,766	450.724	386.099
First fixation duration (ms)	9,476	0	1,190,287	688.823	680.875
Length of Saccades (pixel)	957	0	2,182,187	131.886	87.299
Relative angles of saccades (radian)	3.141	-3.138	-219.437	-0.0133	1.843
Absolute angles of saccades (radian)	3.141	0	18,657.103	1.256	1.020
Number of dynamic Areas Of Interest	18	1	17,402	10.071	3.342
Number of fixations in an AOI	4	1	19,230	1.052	0.247
First degree entropy	1	0	1,618.498	0.937	0.190

TABLE II
RESULTS OF THE PERMUTATION TEST.

Gaze features	R Sum Sq	Pr(Prob)
Sum of the relative angles of the scan path	0.95	2e-16 ***
Number of fixations	0.70	2e-16 ***
Standard deviation of the relative angles of the scan path	0.57	0.06209 .
Standard deviation of the lengths of saccades	0.33	1.00000
Duration of the first fixation	0.23	0.11083
Standard deviation of the absolute angles of the scan path	0.21	0.51579
Number of areas of interest	0.18	0.39103
Sum of the lengths of saccades	0.12	0.07154 .
Mean of the relative angles of the scan path	0.08	0.29150
Mean of the absolute angles of the scan path	0.07	0.65455
Mean of the lengths of saccades	0.06	1.00000
Mean of the distances of fixations	0.03	0.82353
Standard deviation of durations of fixations	0.01	1.00000
First order entropy	0.00	0.42029
Sum of durations of fixations	0.00	1.00000
Mean of durations of fixations	0.00	1.00000
Sum of the absolute angles of the scan path	0.00	1.00000
Mean of the numbers of fixations in an area of interest	0.00	1.00000
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1		
19 observations deleted due to missingness		

TABLE III
RESULTS WHEN CONSIDERING THE EASE OF MEMORIZATION OF EACH CATEGORY.

User	Ease of memorization	Memorization ratio	Nb features p-value<0.05	Nb features p-value<0.01
1	3.76	0.36	0	0
2	3.25	0.22	5	1
3	4.85	0.37	1	0
...
12	2.98	0.13	2	0
15	3.75	0.18	3	0
16	4.52	0.40	3	2
...
21	3.81	0.38	2	2
24	5.05	0.45	2	1

subscribe to a MOOC finish it. The initial interest for this course is not questionable, since all users deliberately chose to subscribe to it, after reading the description and objectives. One factor which could explain these withdrawals is the length and the inappropriate content of the course. These parameters often remain the same regardless of the public, i.e. whatever the knowledge or expertise level the users get. Following this assumption, we propose to envisage a recommender system that fits to the memory state and the learning curve of its users, in addition to their preferences. We thus expect to improve the users' satisfaction and the completion rate of courses, by suggesting resources in adequation with memorized notions and by adapting the learning speed to each student. As this

is a work-in-progress, our first step to achieve this long-term goal consists in studying if one or several of the gaze features are correlated with the memorization of items. We conducted a study with 24 users. All the participants were linked to the academic world (students, teachers, or researchers) and were used to make their memory work. Results show that there was a strong correlation between some gaze features (number of fixations, sum of the relative angles of the scan path) and the fact to memorize items. Thus, these features will be good predictors of what have been learned by students, if we try to analyze in real time their gaze data during a memory test.

These are very promising results, but many future works are required to reach our long-term goal. First, we aim at

conducting additional user studies with different populations, and different tasks based on the reading of texts, charts and multiple sources of information, to confirm these preliminary results according to which there are links between gaze behaviors and working memory. Second, we will extend our protocol to long-term memory. Third, we also aim to go further in the matching between the gaze data and the internal user processes by integrating other models such as the ACT-R [30], in addition to the Atkinson-Shiffrin model used in this paper [5]. The latter has recently been shown to work out well in several learning scenarios [31], [32], [33]. Finally, we will provide machine learning techniques to predict the memory state of users from their gaze data, and recommend items in accordance with their memory.

REFERENCES

- [1] W. Lu and Y. Jia, "Inferring User Preference in Good Abandonment from Eye Movements," in *Web-Age Information Management*, ser. Lecture Notes in Computer Science, J. Li and Y. Sun, Eds., 2015, no. 9098, pp. 457–460.
- [2] L. Chen and P. Pu, "Eye-tracking study of user behavior in recommender interfaces," in *User Modeling, Adaptation, and Personalization*, 2010, pp. 375–380.
- [3] D. Bondareva, C. Conati, R. Feyzi-Behnagh, J. M. Harley, R. Azevedo, and F. Bouchet, "Inferring learning from gaze data during interaction with an environment to support self-regulated learning," in *Artificial Intelligence in Education*, 2013, pp. 229–238.
- [4] D. E. Hannula, C. L. Baym, D. E. Warren, and N. J. Cohen, "The Eyes Know: Eye Movements as a Veridical Index of Memory," *Psychological Science*, vol. 23, no. 3, pp. 278–287, 2012.
- [5] R. M. Shiffrin and R. C. Atkinson, "Storage and retrieval processes in long-term memory," *Psychological Review*, vol. 76, no. 2, p. 179, 1969.
- [6] A. Baddeley, "The episodic buffer: a new component of working memory?" *Trends in cognitive sciences*, vol. 4, no. 11, pp. 417–423, 2000.
- [7] G. A. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information," *Psychological Review*, vol. 63, no. 2, pp. 81–97, 1956.
- [8] F. Gobet and H. A. Simon, "Expert Chess Memory: Revisiting the Chunking Hypothesis," *Memory*, vol. 6, no. 3, pp. 225–255, May 1998.
- [9] D. Jonassen and M. Tessmer, "An outcomes-based taxonomy for the design, evaluation, and research of instructional systems," *Training Research Journal*, vol. 2, no. 1996, p. 97, 1996.
- [10] K. Allan and M. D. Rugg, "An event-related potential study of explicit memory on tests of cued recall and recognition," *Neuropsychologia*, vol. 35, no. 4, pp. 387–397, 1997.
- [11] M. D. Rugg, P. C. Fletcher, K. Allan, C. D. Frith, R. S. J. Frackowiak, and R. J. Dolan, "Neural correlates of memory retrieval during recognition memory and cued recall," *Neuroimage*, vol. 8, no. 3, pp. 262–273, 1998.
- [12] R. B. Clariana and D. Lee, "The effects of recognition and recall study tasks with feedback in a computer-based vocabulary lesson," *Educational Technology Research and Development*, vol. 49, no. 3, pp. 23–36, Sep. 2001.
- [13] J. N. Sari, R. Ferdiana, P. I. Santosa, and L. E. Nugroho, "An eye tracking study: exploration customer behavior on web design," 2015, pp. 69–72.
- [14] C. Kanan, N. A. Ray, D. N. F. Bseiso, J. H. Hsiao, and G. W. Cottrell, "Predicting an Observer's Task Using Multi-fixation Pattern Analysis," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, ser. ETRA '14, NY, USA, 2014, pp. 287–290.
- [15] A. L. Yarbus, *Eye movements during perception of complex objects*. Springer, 1967.
- [16] D. E. Hannula, "Worth a glance: using eye movements to investigate the cognitive neuroscience of memory," *Frontiers in Human Neuroscience*, vol. 4, 2010.
- [17] Z. Bylinskii, P. Isola, C. Bainbridge, A. Torralba, and A. Oliva, "Intrinsic and extrinsic effects on image memorability," *Vision Research*, vol. 116, pp. 165–178, Nov. 2015.
- [18] M. A. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. S. Yeh, D. Borkin, H. Pfister, and A. Oliva, "Beyond Memorability: Visualization Recognition and Recall," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 519–528, Jan. 2016.
- [19] B. Steichen, M. M. Wu, D. Toker, C. Conati, and G. Carenini, "Te, Te, Hi, Hi: Eye gaze sequence analysis for informing user-adaptive information visualizations," in *User Modeling, Adaptation, and Personalization*, 2014, pp. 183–194.
- [20] T. F. Brady, T. Konkle, G. A. Alvarez, and A. Oliva, "Visual long-term memory has a massive storage capacity for object details," *Proceedings of the National Academy of Sciences*, vol. 105, no. 38, pp. 14 325–14 329, 2008.
- [21] T. Konkle, T. F. Brady, G. A. Alvarez, and A. Oliva, "Conceptual distinctiveness supports detailed visual long-term memory for real-world objects," *Journal of Experimental Psychology: General*, vol. 139, no. 3, p. 558, 2010.
- [22] A. M. Maxcey and G. F. Woodman, "Forgetting induced by recognition of visual images," *Visual Cognition*, vol. 22, no. 6, pp. 789–808, 2014.
- [23] M. R. Greene, T. Liu, and J. M. Wolfe, "Reconsidering Yarbus: A failure to predict observers' task from eye movement patterns," *Vision Research*, vol. 62, pp. 1–8, Jun. 2012.
- [24] R. R. Althoff and N. J. Cohen, "Eye-movement-based memory effect: A reprocessing effect in face perception," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 25, no. 4, pp. 997–1010, Jul. 1999.
- [25] T. Karsenti and F. Larose, *Les TIC... Au coeur des pédagogies universitaires : diversité des enjeux pédagogiques et administratifs*. Presses de l'Université du Québec (PUQ), 2001.
- [26] J. Audran and A. Daele, "La socialisation des enseignants au sein des communautés virtuelles : contribution à une compréhension du rapport à la communauté," *Journal of Distance Education*, 2009.
- [27] L. Dawley, *The Tools for Successful Online Teaching*. Idea Group Inc (ICI), 2007.
- [28] A. Boyer, "Les universités numériques thématiques : Bilan," *Rubrique de la Revue STICEF*, vol. 18, 2012.
- [29] A. Roussanally, "Bringing personalized recommendations of open educational resources into the student private digital learning environment," in *E-Learning conference and celebration of the 30th anniversary of the Charter Saar Lor Lux Rhineland-Palatinate Wallony*, 2014.
- [30] M. D. Fleetwood and M. D. Byrne, "Modeling the visual search of displays: A revised act-r model of icon search based on eye-tracking data," *Hum.-Comput. Interact.*, vol. 21, no. 2, pp. 153–197, May 2008.
- [31] P. Seitlinger and T. Ley, "Reconceptualizing imitation in social tagging: a reflective search model of human web interaction," in *WebSci.* ACM, 2016, pp. 146–155.
- [32] P. Seitlinger, D. Kowald, C. Trattner, and T. Ley, "Recommending tags with a model of human categorization," in *CIKM.* ACM, 2013, pp. 2381–2386.
- [33] C. Trattner, D. Kowald, P. Seitlinger, T. Ley, and S. Kopeinik, "Modeling activation processes in human memory to predict the use of tags in social bookmarking systems," *The Journal of Web Science*, vol. 2, no. 1, pp. 1–16, 2016.